Spatial variability patterns of some Vertisol properties at a field scale using standardized data

Humberto Millán , Ana M. Tarquís , Luís D. Pérez , Juan Mato , Mario González-Posada

A B S T R A C T

Spatial variability of Vertisol properties is relevant for identifying those zones with physical degradation. In this sense, one has to face the problem of identifying the origin and distribution of spatial variability patterns. The objectives of the present work were (i) to quantify the spatial structure of different physical properties collected from a Vertisol, (ii) to search for potential correlations between different spatial patterns and (iii) to identify relevant components through multivariate spatial analysis. The study was conducted on a Vertisol (Typic Hapludert) dedicated to sugarcane (*Saccharum officinarum* L.) production during the last sixty years. We used six soil properties collected from a squared grid (225 points) (penetrometer resistance (PR), total porosity, fragmentation dimension (D_f), vertical electrical conductivity (ECv), horizontal electrical conductivity (ECh) and soil water content (WC)). All the original data sets were *z*-transformed before geostatistical analysis. Three different types of semivariogram models were necessary for fitting individual experimental semivariograms. This suggests the different natures of spatial variability patterns. Soil water content rendered the largest nugget effect ($C_0 = 0.933$) while soil total porosity showed the largest range of spatial correlation ($A = 43.92$ m). The bivariate geostatistical analysis also rendered significant cross-semivariance between different paired soil properties. However, four different semivariogram models were required in that case. This indicates an underlying co-regionalization between different soil properties, which is of interest for delineating management zones within sugarcane fields. Cross-semivariograms showed larger correlation ranges than individual, univariate, semivariograms ($A \geq 29$ m). All the findings were supported by multivariate spatial analysis, which showed the influence of soil tillage operations, harvesting machinery and irrigation water distribution on the status of the investigated area.

1. Introduction

Soil attributes usually present a high degree of spatial variation due to a combination of physical, chemical, biological or climatic processes operating at different scales (Goovaerts, 1998; Santra et al., 2008). The quantification and interpretation of such spatial variability is a key issue for site-specific soil management (Brouder et al., 2001). Geostatistics has been the main methodological tool for implementing precision agriculture using field data collected at different spatial resolutions (Burgess and Webster, 1980; Webster, 2008). Many workers have used Geostatistics for comparing the spatial variability between different soil properties and search for corresponding statistical relationships. Castrignanó and Stelluti (1999) used fractal geometry and Geostatistics for describing clay soil aggregation, while Kiliç et al. (2004) related the spatial variability of penetration resistance to some other soil physical properties (e.g. bulk density, clay, sand and water content). García et al. (2008) also studied the spatial variability of soil properties controlling soil resistance to erosion.

Even though all the aforementioned works make significant contributions to the body of knowledge on spatial statistics and its applications, some other key points need to be addressed for conducting precise comparisons between soil properties using geostatistical parameters. First, different soil properties have different open ranges of variation in terms of numerical values (e.g. electrical and hydraulic conductivity, penetration resistance and many others). On the contrary, many soil characteristics present bounded values (e.g. total porosity, gravimetric or volumetric water content, clay, silt and sand content can take a maximum value equal to unity or 100%). Second, all soil variables use different types of scales. Thus, one requires removing those inconveniences before any analysis. That is equivalent to set all the investigated variables within the same scaling framework and to avoid, to some extent, any biased interpretation. To our knowledge, Eghball and

Varvel (1997) have taken into account that situation. Those authors used standardized data for characterizing spatial and temporal yield variability of crop sequences using Geostatistics and fractal analysis. Kravchenko et al. (1999), working with chemical properties, standardized the variables for conducting multifractal analysis of soil spatial variability. Grunwald et al. (2001) also standardized the variables before fuzzy clustering. Although the geostatistical literature is ample, that procedure has been less considered in most investigations. In addition of setting all the variables within the same framework, one has to face the problem of identifying the origin and distribution of spatial variability patterns. The objectives of the present work were (i) to quantify the spatial structure of different physical properties collected from a Vertisol, (ii) to search for potential correlations between different spatial patterns and (iii) to identify relevant components through multivariate spatial analysis.

2. Materials and methods

2.1. Study site

The field measurements and soil samples were obtained in Bayamo, Granma Province, Cuba (20°22′N, 76°38′W, approximately 50 m asl) (Fig. 1). The soil is classified as a Vertisol, Typic Hapludert, according to Soil Survey Staff (2003) standards. Some physical and chemical characteristics of the studied field are reported in Pérez et al. (2010). The study site has been under sugarcane (*Saccharum officinarum* sp.) production during the last 60 years. The present investigation was conducted in March–April/2009 after sugarcane harvest.

2.2. Sampling design and data collection

Each data set was collected from 225 points of a squared grid ($15 \times 15 = 225$ points) with sampling interval $L = 10$ m. Fig. 2 presents the spatial position for the particular case of penetrometer resistance readings. All soil samples and instrument readings were obtained within the 0–10 cm soil layer. Penetrometer resistance (PR), total porosity, fragmentation dimension (D_f), vertical electrical conductivity (ECv), horizontal electrical conductivity (ECh) and soil water content (WC) values were estimated from each selected sampling point. Soil resistance measurements were carried out using an electronic penetrometer (FIELDSCOUT™ SC900 Soil Compaction Meter, Spectrum Technologies Inc., Illinois). In this particular case, we considered only those readings obtained for the 7.5–10 cm depth range. Undisturbed soil samples were extracted from each point using cylindrical soil cores (3.6 inner diameter, 10 cm height). Each cylindrical soil core was oven-dried at 105 °C during 24 h for bulk density and gravimetric water content estimations. No corrections were performed for soil volume decrease after oven drying. The aforementioned soil properties were computed after weighing each oven-dried sample. In particular, total porosity was estimated as:

$$\phi = 1 - \frac{\rho_b}{\rho_p} \tag{1}$$

where ϕ is the total soil porosity, ρ_b is the soil bulk density and ρ_p is the particle density, in this case, $\rho_p = 2.65$ Mg m^{-3}.

Both, ECh and ECv were estimated using an electromagnetic induction probe (EM-38 model, frequency = 13.2 kHz, range 0–1.0 S m^{-1}, accuracy = 0.01 S m^{-1}). We recall here that horizontal

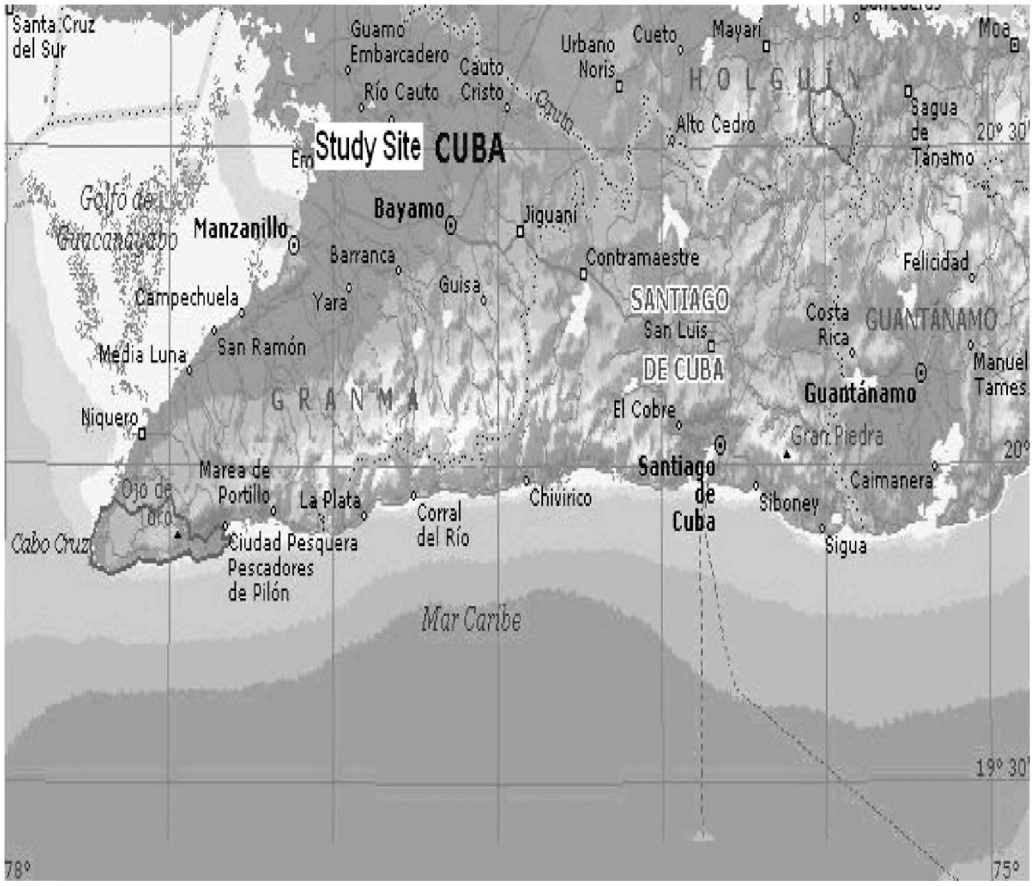

Fig. 1. Physical map of the study area.

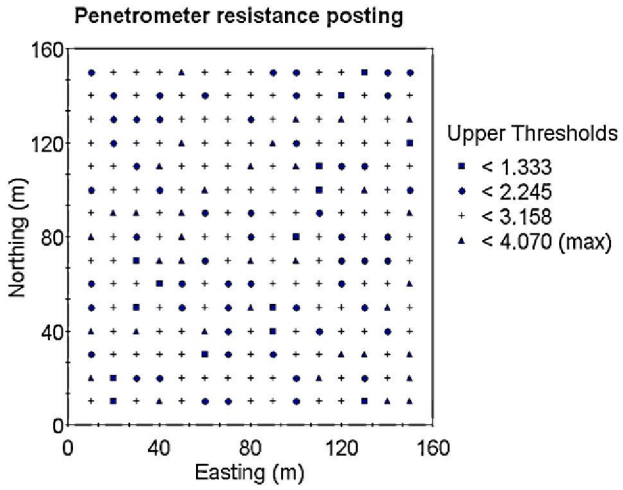

Fig. 2. Squared grid showing the sampling points (legend corresponds to penetrometer resistance readings).

(e.g. ECh) and vertical (ECv) terms refer to instrument dipole position. McNeill (1986) and Slavich (1990) have presented the basic rationale on both terms. In reality, both types of measurements can span approximately 2.0 m in depth (Millán et al., 2009).

A flat shovel was used for extracting undisturbed soil clods from each point of the squared grid. At the laboratory, the clods (mass ≈232 ± 41 g) were air dried on trays during 15 days. After that, each soil block was fragmented under the impact of a led plate (mass = 5 kg) falling down from a height of h = 1 m. This procedure imparts approximately 50 J of potential energy to each soil clod (Millán, 2004a,b). After the fragmentation process, all the soil material was placed over the uppermost sieve of a nest and sieved during t = 30 s as suggested by Díaz-Sorita et al. (2002). Nine soil fragment classes were used: <0.125, 0.125–0.250, 0.250–0.5, 0.5–1.0, 1.0–2.0, 2.0–3.0, 3.0–5.0, 5.0–7.0 and 7.0–10.0 mm.

We computed the cumulative mass of soil fragments retained under each upper class limit. Each cumulative mass-size distribution was fitted to the Pore Solid Fractal (PSF) model (Perrier and Bird, 2002; Millán, 2004a,b):

$$m(x \le x_i) = Bx_i^{3-D_f} \tag{2}$$

where m is the cumulative mass of fragments with size $x \le x_i$, x_i = 0.125, 0.250, 0.5, 1.0, 2.0, 3.0, 5.0, 7.0 and 10.0 mm, D_f is the fragmentation dimension and B is a parameter accounting for the unfragmented soil material (e.g. probability of fragmentation). Within the context of the present work, D_f represents a measure of soil fragmentation as influenced by external forces (e.g. machinery traffic or tillage operations). A log–log transformation of Eq. (2) allows one to estimate D_f from the slope of the linear regression equation:

$$\log[m(x \le x_i)] = \log(B) + (3 - D_f)\log(x_i) \tag{3}$$

In this case, we did not consider any analysis of the B parameter.

2.3. Exploratory data analysis

Each spatial soil property distribution was characterized with its descriptive statistics. Kolmogorov–Smirnov tests of normality ($p < 0.05$ indicating significant differences from the normal distribution) were also performed on each data set. Prior to any analysis, each data set was standardized to a zero mean and unit variance:

$$Y_{st.} = \frac{(Y - \mu)}{s_Y} \tag{4}$$

where $Y_{st.}$ is the new standardized value, Y is the estimated value of the soil property, μ is the mean and s_Y is the standard deviation. Here one has to assume that μ is constant, at least locally (Webster, 2008). The application of Eq. (4) permits one to compare each variable within the same scaling range (Eghball and Power, 1995). Kolmogorov–Smirnov tests of normality, descriptive statistics, data standardization and linear regression analysis were performed using STATISTICA™ Software Package (Stat. Soft. Inc., 2003).

2.4. Linear geostatistical analysis

Each standardized soil property data set was analyzed using geostatistical methods. Experimental semivariograms were calculated for each standardized variable as follows (Journel and Huijbregts, 1978):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Y_{st.}(x_i + h) - Y_{st.}(x_i)]^2 \tag{5}$$

where $\gamma(h)$ is the value of experimental semivariance at distance lag h, $N(h)$ is the number of data pairs separated by the lag h, $Y_{st.}(x_i)$ and $Y_{st.}(x_i + h)$ are standardized data values at two points separated by lag distance h.

We used each experimental semivariogram for constructing contour maps of the interpolated variable. In such a case, block kriging was considered as the most suitable working tool:

$$Y^*_{st.}(x_0) = \sum_{i=1}^{n} \lambda_i Y_{st.}(x_i) \tag{6}$$

where $Y^*_{st.}$ is the estimated kriged value of $Y_{st.}$ at the point x_0 and λ_i refers to weighing factors such that:

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{7}$$

Accuracy of the kriging maps was evaluated through a cross-validation process using the Mean Absolute Error (MAE hereafter) (Voltz and Webster, 1990) and goodness-of-prediction statistics (G hereafter) (Agterberg, 1984).

The MAE statistics refers to the residual sum:

$$\mathrm{MAE} = \frac{1}{N} \sum_{i=1}^{N} [|Y_{st.}(x_i) - \hat{Y}_{st.}(x_i)|] \tag{8}$$

where $\hat{Y}_{st.}(x_i)$ is the predicted standardized value at point i.

The G parameter compares, to some extend, kriging and sample mean as potential predictors:

$$G = \left(1 - \frac{\sum_{i=1}^{N} [Y_{st.}(x_i) - \hat{Y}_{st.}(x_i)]^2}{\sum_{i=1}^{N} [Y_{st.}(x_i) - Y_{st.}(\mathrm{mean})]^2}\right) \times 100 \tag{9}$$

where $Y_{st.}$(mean) is the mean of the sample. Note that, within the context of the present work, $Y_{st.}$(mean) = 0 as we used standardized data.

Even without using standardized data, one can interpret three different situations from Eq. (9).

(i) G = 0 indicates that kriged values and sample mean can be used as predictors,
(ii) $G < 0$ suggests that sample mean is a more reliable predictor as compared with kriging,
(iii) $G > 0$ indicates that kriging is more effective than sample mean for making consistent predictions at unsampled sites.

The Geostatistical software Vesper 1.6 (unregistered version) (Minasny et al., 2002) was used for fitting the appropriate semivariogram model to each standardized data set. The following computational parameters were considered: maximum distance = 160 m, number of lags = 30, lag tolerance = 50%. We retained the software default values for the number of lags and lag tolerance (%). Due to the selected lag tolerance, a pair of points separated by a distance of 10 m (e.g. present work) has a tolerance of 10 ± 5 m. The criteria for model selection were based on manual fitting, Akaike Information Criterion (AIC) and Root Mean Squared Error (RMSE).

In particular, the AIC statistics accounts for the balance between data parsimony and goodness of fit (Akaike, 1973):

$$\mathrm{AIC} = n \ln\left(\frac{\mathrm{RSS}}{n}\right) + 2p \tag{10}$$

where n is the number of data points (n = 30 lags), RSS is the residual sum of squares and p is the number of model parameters. A smaller value of the AIC statistics is associated with higher goodness of fit of the selected model.

In addition, we performed exploratory cross-semivariance analyses using fifteen different pairs of variables. The joint spatial variability of two soil properties, as represented by $Y^i_{\mathrm{st.}}$ and $Y^j_{\mathrm{st.}}$, can be represented by the experimental cross-semivariogram (Goovaerts, 1999):

$$\hat{\gamma}_{i,j} = \frac{1}{2N(h)} \sum_{\beta=1}^{N(h)} [Y^i_{\mathrm{st.}}(x_\beta) - Y^i_{\mathrm{st.}}(x_\beta + h)] \times [Y^j_{\mathrm{st.}}(x_\beta) - Y^j_{\mathrm{st.}}(x_\beta + h)] \tag{11}$$

where β = 1, 2, 3, …, n locations.

The GS + Geostatistical Software Package (Gamma Design Software, 2001) was used for performing cross-semivariance (angular tolerance of 22.5°), kriging and cross-validation analyses. We selected 4 m × 4 m local grids and16 neighbors within a radius equal to the range of the semivariogram for kriging interpolation.

2.5. Exploratory multivariate spatial analysis

Principal Component Analysis (PCA hereafter) is a useful tool for extracting leading patterns from high-dimensional datasets. This uses an orthogonal transformation for converting a set of dependent variables into a reduced number of independent variables. Those new variables are called principal components. In this case, the six investigated soil variables were grouped into two principal components (Varimax Rotation) under the criterion of selecting eigenvalues >1. Each PC can be considered as a new soil variable. PC score sets were used for geostatistical analysis and kriging interpolation. This multivariate spatial analysis was carried out using STATISTICA™ Software Package (Stat. Soft. Inc., 2003). Thus, the main purpose of geostatistical analysis of the first two PCs was to integrate correlated soil properties into kriging maps.

3. Results and discussion

3.1. Exploratory data analysis and linear statistics

Table 1 shows the main descriptive statistics for each considered data set before standardization. In terms of the coefficient of variation, the fragmentation dimension was the less variable soil parameter (CV = 7.7%) while medium variability can be considered for the rest of the variables (CV ranging from 11.5 to 27.7%). The mean value of penetration resistance was near to the limit 2.5 MPa, which is considered as critical for root growth (Carrara et al., 2007). Note also that penetration resistance presented the largest range of data variability (range = 3.649 MPa) while the fragmentation dimension showed the second largest range of variation (range = 0.851). This indicates, to some extent, that neither the coefficient of variation nor rough data sets are appropriate for evaluating the spatial patterns of soil variables at a field scale. Mean values of ECv and ECh were larger if one considers that sugarcane is a herbaceous crop with a smaller salt tolerance threshold (e.g. ECe of the saturated soil extract ≈0.170 S m^{-1} for sugarcane) (Dev and Bajwa, 1972). Thus, one can advance that potential yield of the crop could be compromised.

Fig. 3(a)–(f) illustrates that three (penetration resistance, ECv and ECh) out of six data sets were normally distributed according to the Kolmogorov–Smirnov test of normality ($p < 0.05$). Neither log-transformation nor square root transformation improved significantly the other three distributions (total porosity, D_f and water content). Since skewness values ranged approximately from −1.0 to +1.0 for the other three variables (Table 1), no other transformations were considered for geostatistical analysis.

We found a significant linear relationship between D_f and total porosity before data standardization:

$$D_f = 2.4661(0.08) + 0.626(0.21)\phi \quad R = 0.213,\ n = 225,\ p < 0.05 \tag{12}$$

It is evident from Eq. (12) that the extreme case $\phi \to 1$ suggests $D_f \to 3$. Physically, D_f is an indirect measure of pore space and failure planes complex geometry while total porosity accounts for the number and size of micro-, meso-, macropores, fissures and microcracks (Millán, 2004a,b). We also found a linear relationship between ECv and total porosity:

$$\mathrm{ECv} = 0.183(0.02) + 0.111(0.04)\phi \quad R = 0.175,\ n = 225,\ p < 0.05 \tag{13}$$

That is, the larger the total soil porosity, the larger the soil volume occupied with soil solution.

Table 1 also shows that ECv values were larger than ECh ones. From an environmental point of view that could be problematic due to the possibility of a contaminated groundwater table. However, we have no data on the chemistry of soil solution composition. Furthermore, we also found a linear dependence

Table 1
Descriptive statistics corresponding to each selected variable.

Variable	Mean	Minimum	Maximum	Range	Skewness	Kurtosis	CV (%)
PR (MPa)	2.493	0.421	4.070	3.649	−0.314	0.411	27.7
Porosity (m^3 m^{-3})	0.419	0.245	0.671	0.426	0.518	1.345	15.6
D_f	2.729	2.10	2.951	0.851	−0.850	0.503	7.7
ECv (S m^{-1})	0.230	0.133	0.354	0.221	0.608	0.458	18.1
ECh (S m^{-1})	0.185	0.118	0.287	0.169	0.329	−0.241	19.9
WC (kg kg^{-1})	0.201	0.171	0.301	0.130	0.991	1.764	11.5

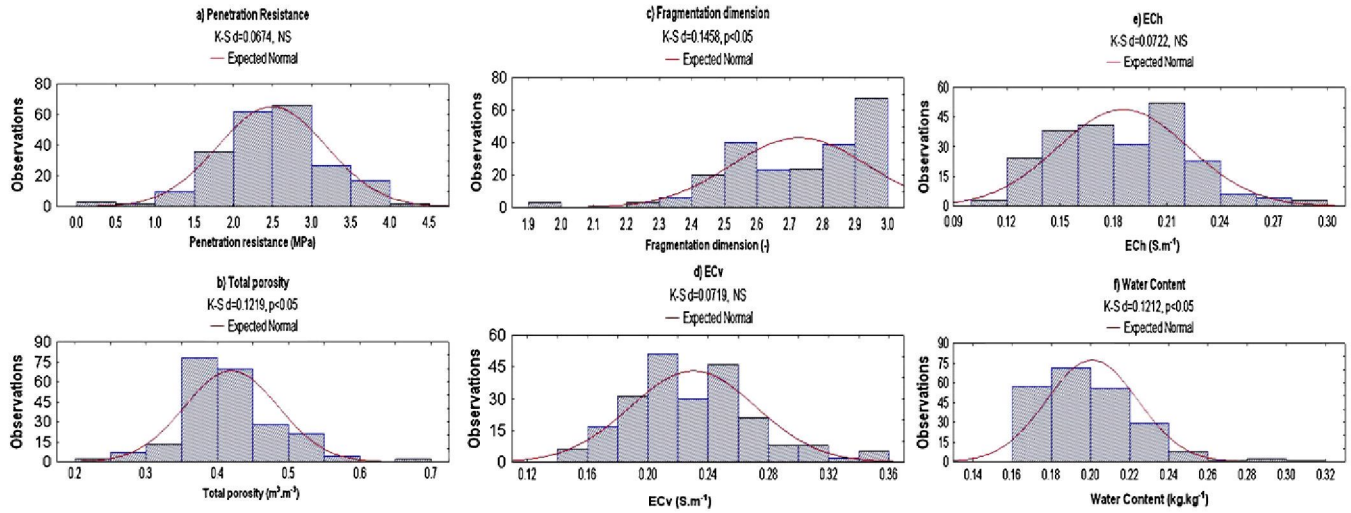

Fig. 3. Histograms of untransformed soil properties.

between both variables:

$$\mathrm{ECh} = 0.014(\mathrm{NS}) + 0.744(0.03)\mathrm{ECv} \quad R = 0.842,\ n = 255,\ p < 0.05 \tag{14}$$

That is an expected finding as liquid phase moves upward mainly under the effect of capillary force.

3.2. Geostatistical analysis

Fig. 4(a)–(f) presents each fitted semivariogram model. The minimum number of pairs of points for calculating the semivariance at each lag was $N = 124$ while the maximum number was $N = 2630$. At a first approximation, one can use Cambardella et al. (1994) rationale for classifying the spatial pattern corresponding to each standardized variable. Those authors stated the ratio nugget semivariance/total semivariance as a first indicator of spatial dependence. The soil variable is strongly spatially dependent if the ratio $\leq$25%; if that ratio ranges from 26% to 75%, the soil variable is considered as moderately spatially dependent, while a ratio ranging from 75% to approximately 95% classifies the variable as weakly spatially dependent. One important advantage of standardizing the data set is that total semivariance is always equal to unity (note that $C_0 + C \approx 1$) (Table 2). This means that the intercept with the y-axis (nugget semivariance) corresponds exactly to the percentage of the total semivariance accounted for the nugget effect. Under that rationale, penetrometer resistance (nugget semivariance = 79.8%), ECv (nugget semivariance = 85.9%) and gravimetric water content (nugget semivariance = 93.3%) presented weak spatial dependence. On the other hand, total porosity

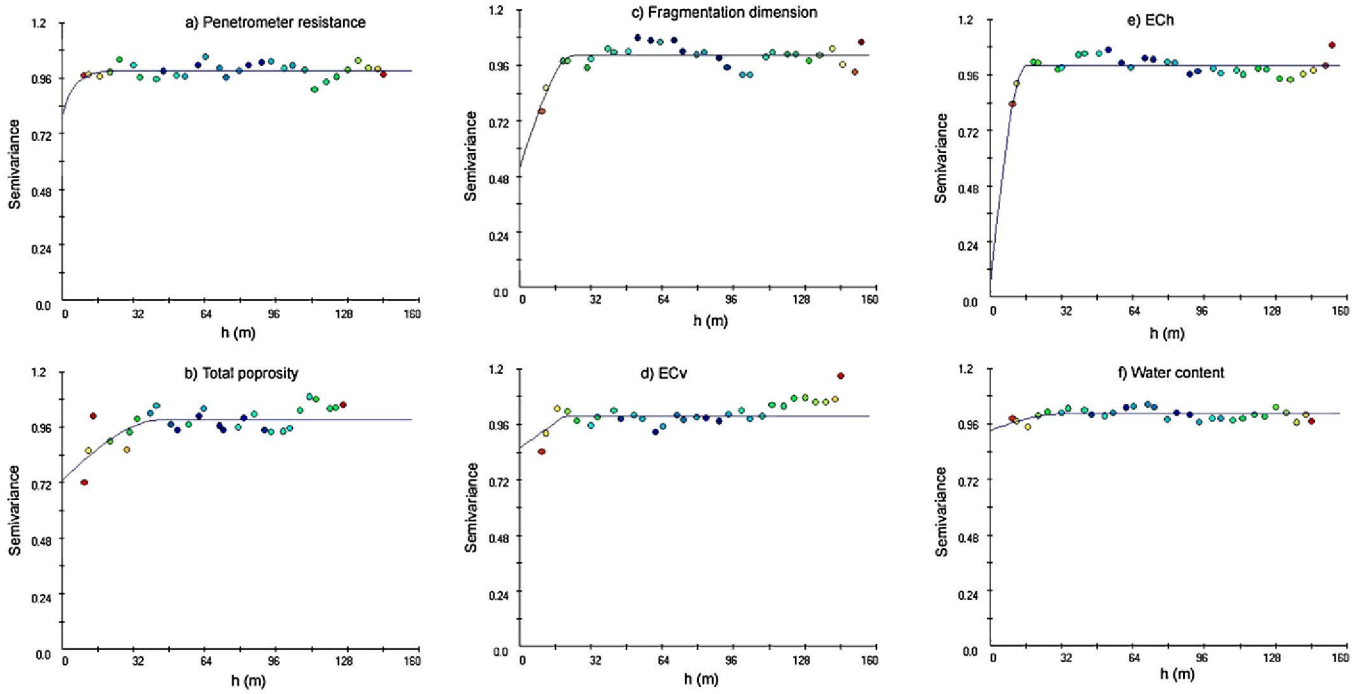

Fig. 4. Univariate semivariogram for each standardized soil property (lag color represents the number of pair of points used for calculating the semivariance at each lag. The number of pair of points increases from pink to blue color). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Parameters derived from the geostatistical analysis (standardized variables).

Variable	Model	C_0	C^a	A (m)	AIC	RMSE
PR	Exponential	0.798	0.196	13.16	−100.8	0.031
Porosity	Spherical	0.731	0.260	43.92	−64.08	0.056
D_f	Spherical	0.514	0.492	24.71	−82.29	0.042
ECv	Linear with sill	0.859	0.139	20.44	−66.00	0.055
ECh	Spherical	0.056	0.945	15.81	−92.33	0.035
WC	Spherical	0.933	0.071	32.54	−121.9	0.022

[a] Note that $C_0 + C \approx 1$ due to data standardization. C_0 is the nugget semivariance, C is the structural semivariance and A is the correlation range.

(nugget semivariance = 73.1%) and fragmentation dimension (nugget semivariance = 51.4%) showed moderate spatial dependence. If one takes into account the intercept with the *y*-axis, then ECh was the more spatially dependent pattern (nugget semivariance = 5.6%). That particular value needs careful interpretation even though the manual fitting corresponded to the smaller possible AIC and RMSE values. Almost all the investigated patterns showed higher nugget effect, which accounts for short-scale spatial variability or measurement errors. Note also that all the considered soil variables have direct or indirect connection to soil pore system. Thus, regardless of the influence of data acquisition errors, the nugget semivariance also reflects the short-scale complexity of soil pore system.

Each spatial pattern showed different ranges of spatial dependence. Total porosity, gravimetric water content, fragmentation dimension and ECv produced ranges of correlation larger than 20 m. However, we required three different semivariogram models for describing the standardized semivariograms (e.g. exponential, spherical and linear with sill models) (Table 2). Each semivariogram model sheds light on the cause–effect mechanism of the spatial process. For example, an exponential model can be associated with a first order autoregressive (Markov process) or a Poisson process (Kuzyakova et al., 2001). Physically, a Poisson process is always related to external sources. That is consistent with soil compaction, which is usually associated to machinery traffic or soil tillage operations. In agricultural practice, soil compaction can be estimated from penetration resistance measurements or bulk density data.

Semivariogram spherical models represent moving average of randomized processes (Kuzyakova et al., 2001). It is interesting that both, D_f and total porosity spatial patterns were described by a spherical model (Table 2). It is known the high complexity of soil pore, crack and fissure organization and its influence on soil fragmentation process. Fractal or multifractal models usually capture the dynamics of such a complex fragmentation (Perfect et al., 2002; Martín and Montero, 2002).

Fig. 5(a)–(f) presents the significant cross-semivariograms while Table 3 shows the geostatistical parameters corresponding to the selected model. We point out here that our main goal was to identify meaningful spatial relationships between the variables investigated instead of cokriging maps. Six out of fifteen variable combinations rendered significant co-regionalization between paired soil property patterns. Each paired spatial pattern was significantly interrelated. That is, an increase in $\Delta Y^i_{st.} = Y^i_{st.}(x_\beta) - Y^i_{st.}(x_\beta + h)$ could correspond to an increase or decrease in $\Delta Y^j_{st.} = Y^j_{st.}(x_\beta) - Y^j_{st.}(x_\beta + h)$ (Goovaerts, 1999). We add that only the differences are important and not the attribute value by itself. For example, penetrometer resistance/fragmentation dimension, penetrometer resistance/total porosity and penetrometer resistance/water content differences should be negatively correlated. On the contrary, soil total porosity/fragmentation dimension and ECv/ECh differences might be positively correlated. From a cause–effect perspective, the spatial variability of soil compaction (e.g. machinery traffic) influences the spatial variability of both, soil hydraulic properties (e.g. total porosity) and the result of tillage operations (soil fragmentation). Such information could be valuable for soil users. In particular, the joint spatial variability of total porosity versus D_f patterns looks like a pseudo-cyclic structure (Fig. 5(e)). Thus, we used a spherical model since a periodic model is difficult to apply. In this case, 2-dimensional data represents a relatively uniform surface with patches with contrasting soil properties such that semivariance tends to

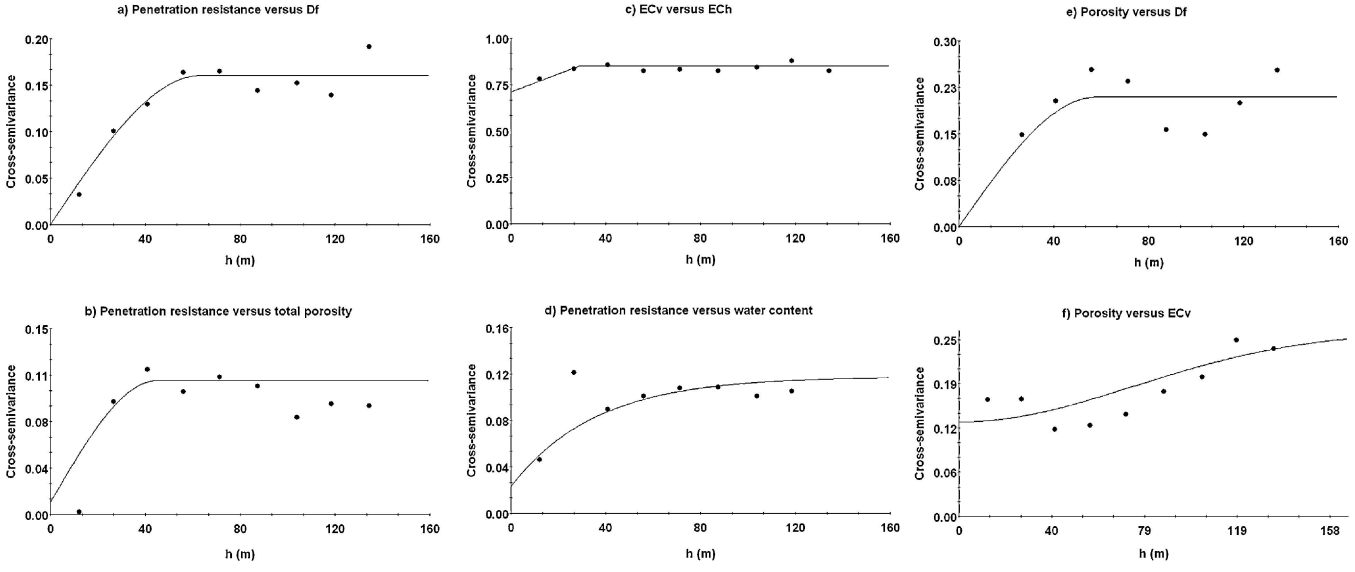

Fig. 5. Bivariate cross-semivariogram for spatial pattern interactions.

Table 3
Parameters derived from the significant cross-semivariance analysis (standardized variables).

Variable	Model	C_0	C	A (m)	R^a	RSS
PR × D_f	Spherical	0.000	0.159	63.2	0.941	0.002
PR × porosity	Spherical	0.011	0.098	44.9	0.894	0.003
ECv × ECh	Linear with sill	0.711	0.140	28.9	0.851	0.003
PR × WC	Exponential	0.023	0.095	59.8	0.801	0.005
Porosity × D_f	Spherical	0.000	0.209	57.6	0.861	0.002
Porosity × ECv	Gaussian	0.132	0.138	105.5	0.821	0.006

[a] R is the correlation coefficient.

oscillate. That type of quasi-periodic structure or spot effect (Krasilnikov, 2008) has been previously observed from experimental semivariograms. For instance, Webster (1977) found quasi-periodic patterns in some soil properties (e.g. pH, electrical conductivity) collected along a 1-km transect. It is possible that cyclic gilgai microrelief and blocky structure of Vertisols generate such pseudo-periodic (incomplete oscillation cycle) patterns. This also agrees with Webster's (1977) considerations. However, we found such a quasi-cyclic structure from a cross-semivariogram. In general, the nugget effect was smaller for the cross-semivariance analysis (Table 3) as compared with individual semivariograms (Table 2). This fact suggests useful information as the interaction between different soil properties might be more important than experimental errors regarding short-range spatial variability (e.g. $h \rightarrow 0$). In particular, there is little information on the spatial variability of soil fragmentation dimension at a field scale. We think however that it could be a useful soil parameter even though its estimation represents a time-consuming work. In this case, a potential utility of cross-semivariance is that one can predict the geostatistical properties of soil fragmentation from others, easily to collect, soil variables (e.g. penetrometer resistance).

Table 5 shows the soil variable loadings for PC1 and PC2 (eigenvalues >1) based on correlation coefficients. PC1 accounted for the correlation between ECv and ECh (31.47% of the total variance). This finding is interesting as it sheds light on the influence of irrigation water quality on the spatial pattern of electrical conductivity. Here, we take into account that the study field is far away from coastal areas and the main source of potential salinity is through irrigation. Thus, it takes one to the need of controlling the salt level of irrigation water. PC2 grouped four soil variables (PR, total porosity, D_f and WC) which accounted for 24.75% of the total variance (Table 5). Within the context of the present work, the PC2 is of utmost importance as it accounted for the interplay between those variables associated to harvest machinery traffic (e.g. PR) and soil tillage operations (e.g. D_f, WC and total porosity). Table 6 shows that spherical semivariogram models fitted reasonably well both PC experimental semivariograms. We found the best fitting for 120 m as the maximal distance instead of 150 m for individual fitting. Note that nugget effect was smaller when compared with individual semivariograms. This supports our assumption that short-scale variability can be reduced when one considers the influence of others soil variables. In this case any influence of measurement errors is minimized.

3.3. Kriging maps

Fig. 6(a)–(f) shows the kriging map corresponding to each individual spatial pattern. In particular, we highlighted two zones with contrasting values of predicted soil total porosity and fragmentation dimension (see Fig. 6(b) and (c), respectively). Note

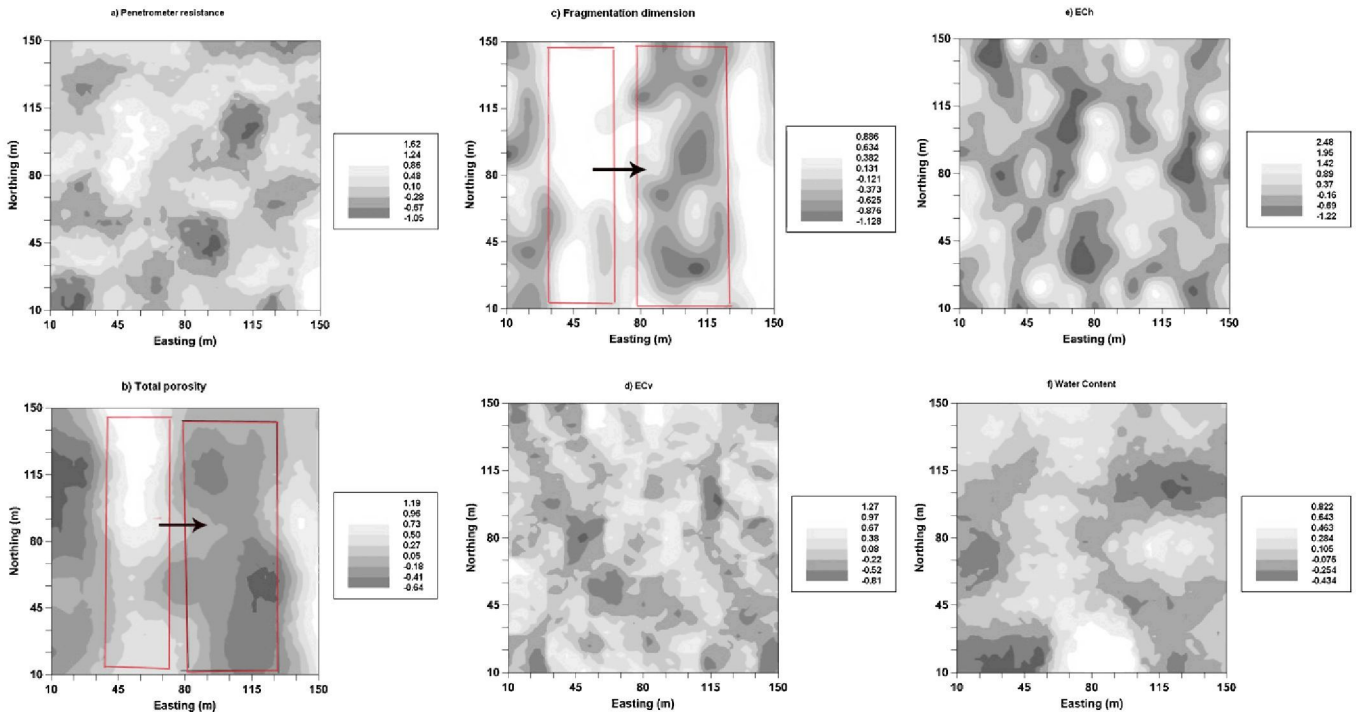

Fig. 6. Kriging maps for each standardized soil property. Rectangles indicate two zones with contrasting soil properties.

Table 4
Cross-validation statistics.

Variable	MAE	G (%)
PR	0.0005	−23
Porosity	0.0015	+18
D_f	0.0019	+11
ECv	0.0105	−10.2
ECh	0.0059	+8
WC	0.0015	+5

that areas of higher/smaller soil total porosity match reasonably well those areas of higher/smaller fragmentation dimension. Both soil attributes are important for exploring the physical quality of the soil. The kriging map of penetrometer resistance predicts smaller values than the mean (e.g. standardized values <0.00) in those patches with the largest values of fragmentation dimension (Fig. 6(a) and (c), respectively). Furthermore, the kriging interpolation predicted that approximately 60% of the field possess penetrometer resistance values larger than the mean (positive values in the bar scale). A practical question in this case is how we can make sure this information is valuable to farmers or soil users. For example, based on Fig. 6(b) and (c) one can detect a complete zone and several patches of potentially lower crop yield. An interesting idea could be to establish different crops within the field for restoring the degraded zones. According to cross-validation statistics (Table 4), mean PR value could be used as a consistent predictor at unsample points instead of kriged values (G = −23%). That is, the PR value at an unsampled point can be set as the mean value of sampled points.

There exists some sort of similarity between the maps corresponding to ECv (Fig. 6(d)) and ECh (Fig. 6(e)) which gives a picture on the salinity pattern at the soil surface and those deeper soil layers. Nevertheless, mean ECv value could be also a better predictor when compared with kriging as G = −10.2% for that variable (Table 4). At the same time, the kriging map of soil moisture suggests a non-uniform distribution within the field (Fig. 6(f)). Kriging interpolation predicted higher values of soil moisture at both extremes of the field (north-east direction) which is due to the influence of irrigation channels. In general, more than 60% of the field showed higher values of predicted soil water content. This field is evenly irrigated in the dry season through an underground pipe system. However, Fig. 6(f) indicates that site-specific irrigation could render better results with important water savings. All those kriging maps can be used for delineating crop management zones in those sugarcane fields. For example, we recommend manual cutting of sugarcane in those patches with smaller values of kriged total porosity (Fig. 6(b)). That could be a time-consuming task but it would render environmental benefits. In general, four out of six G values were greater than zero, which indicates that semivariogram model and kriging are effective tools for spatial prediction. Santra et al. (2008) found similar results for other chemical and physical soil variables. However, the nugget effect for penetrometer resistance and ECv showed larger values (Table 2). According to these findings and Table 3, cokriging maps could be very useful for final decision-making. A large scale study on this topic is currently underway.

Fig. 7(a) and (b) shows the multivariate kriging maps for PC1 and PC2. At a first sight, the same contrasting zones detected from individual semivariograms are preserved. However, one can note continuous zones with positive PC1 scores (lighter gray areas), which indicates a strong interaction between ECh and ECv and the first PC. Here, the downward movement of saline compounds should be the predominant mechanism. The similarity of Fig. 6(b) and (c) to the kriging map corresponding to PC2 (Fig. 7(b)) confirms the presence of two zones with quite different characteristics and

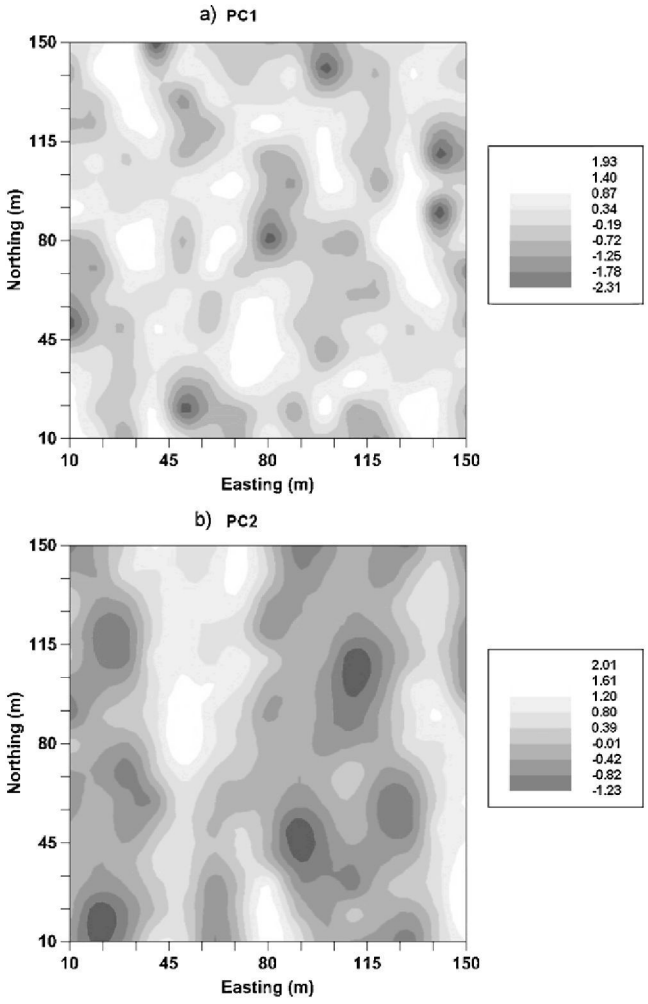

Fig. 7. Multivariate kriging showing the two leading PCs: (a) PC1 and (b) PC2.

the relevance of total porosity and D_f for defining such structures. At this stage, we should stress the important interplay between D_f (loading = 0.679, Table 5) and water content (WC). That is, depending on water content soil clods can fail as brittle fragmentation (dry soil) or plastic failure (wet soil) after tillage operations. Those fragmentation modes determine different qualities of the soil structure (Hadas, 1997; Millán, 2004a). Such a situation can be interpreted from Figs. 6(f) and 7(b) where WC shows contrasting patterns. In general terms, sugarcane producers need to address two fundamental questions: first of all to develop strategies for controlling irrigation water quality and its site-specific distribution within the field and second, to minimize the

Table 5
Results of multivariate spatial analysis.

Soil variable	PC1	PC2
PR	NS*	0.593
Porosity	NS	0.564
D_f	NS	0.679
ECv	−0.953	NS
ECh	−0.944	NS
WC	NS	0.503
Eigenvalue	1.888	1.365
% total variance	31.47	24.75

* Significant correlations, $R > 0.50$

Table 6
Parameters derived from the geostatistical analysis (principal components).

Parameter	PC1	PC2
Model	Spherical	Spherical
Maximal distance (m)	120	120
C_0	0.004	0.251
C	0.985	0.781
A (m)	15.07	29.55
AIC	−101.1	−71.46
RMSE	0.0293	0.0486
R	0.810	0.851

impact of sugarcane harvesting machinery and subsequent tillage operations (Table 6).

4. Conclusions

We investigated the spatial structure of six standardized soil variables collected at a mesoscopic scale (e.g. field scale). It was included the fragmentation dimension as a relevant parameter accounting for the results of tillage operations. Three different semivariogram models were necessary for fitting the spatial data sets. This suggests the different natures of spatial variability. Gravimetric water content rendered the largest nugget effect while soil total porosity showed the largest range of spatial correlation. The bivariate geostatistical analysis also rendered significant cross-semivariance between different paired soil properties. However, four different semivariogram models were required in that case. This indicates an underlying co-regionalization between different soil properties, which is of interest for delineating management zones within agricultural fields. In general, cross-semivariograms showed larger correlation ranges than individual semivariograms. The multivariate spatial analysis (PCA) suggested two principal components accounting for the influence of soil electrical conductivity (PC1) and soil total porosity, gravimetric water content, fragmentation dimension and penetration resistance (PC2). This information could be used for delineating site-specific irrigation, sugarcane harvesting and subsequent soil tillage operations.

Acknowledgments

We want to express our gratitude to several undergraduate and graduate students who enthusiastically participated in the soil sampling. We also want to acknowledge the referees and the editorial board for the time and very useful comments and suggestions that help us to improve the final version of the manuscript.

References

Agterberg, F.P., 1984. Trend surface analysis. In: Gaile, G.L., Willmot, C.J. (Eds.), Spatial Statistics and Models. Reidel, Dordrecht, The Netherlands, pp. 147–171.
Akaike, H., 1973. Information theory and an extension of maximum likelihood principle. In: Petrov, B.N., Csáki, F. (Eds.), Second International Symposium on Information Theory. Akadémia Kiadó, Budapest, pp. 267–281.
Brouder, S., Hofmann, B., Reetz, H.F., 2001. Evaluating spatial variability of soil parameters for input management. Better Crops 85, 8–11.
Burgess, T.M., Webster, R., 1980. Optimal interpolation and isarithmic mapping of soil properties. I: The semivariogram and punctual kriging. J. Soil Sci. 31, 315–333.
Cambardella, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Karlen, D.K., Turco, R.F., Konopka, A.E., 1994. Field-scale variability of soil properties in central Iowa soils. Soil Sci. Soc. Am. J. 58, 1501–1511.
Carrara, M., Castrignanó, A., Comparetti, A., Febo, P., Orlando, S., 2007. Mapping of penetrometer resistance in relation to tractor traffic using multivariate geostatistics. Geoderma 142, 294–307.
Castrignanó, A., Stelluti, M., 1999. Fractal geometry and geostatistics for describing the field variability of soil aggregation. J. Agric. Eng. Res. 73, 13–18.
Dev, G., Bajwa, M.S., 1972. Studies on salt tolerante of sugarcane. Ind. Sugar (Trinidad) 22, 723–726.
Díaz-Sorita, M., Perfect, E., Grove, J.H., 2002. Disruptive methods for assessing soil structure. Soil Tillage Res. 64, 3–22.
Eghball, B., Power, J.F., 1995. Fractal description of temporal yield variability of 10 crops in the United States. Agron. J. 87, 152–156.
Eghball, B., Varvel, G.E., 1997. Fractal analysis of temporal yield variability of crop sequences: implications for site-specific management. Agron. J. 89, 851–855.
Gamma Design Software, 2001. GS+ Geostatistics for the Environmental Sciences, Version 5.1.1. Professional Edition, Plainwell, MI.
García, N.E., Uriostegui, Y., Alvarez, G., Ibáñez, J., Krasilnikov, P., 2008. Spatial distribution of the soil properties controlling soil resistance to erosion at a coffee growing farm in Sierra Sur de Oaxaca. In: Krasilnikov, P.,Carré, F., Montanarella, L. (Eds.),Soil Geography and Geostatistics. Concepts and Applications. JRC Scientific and Technical Reports. Office for Official Publications of the European Communities, Luxembourg, (Chapter 4), pp. 37–45.
Goovaerts, P., 1998. Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. Biol. Fertil. Soil. 27, 315–334.
Goovaerts, P., 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma 89, 1–45.
Grunwald, S., McSweeney, K., Rooney, D.J., Lowery, B., 2001. Soil layer models created with profile cone penetrometer data. Geoderma 103, 181–201.
Hadas, A., 1997. Soil tilth—the desired soil structural state obtained through proper soil fragmentation and reorientation processes. Soil Tillage Res. 43, 7–40.
Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London, United Kingdom.
Kiliç, K., Özgöz, E., Akbaş, F., 2004. Assessment of spatial variability in penetration resistance as related to some soil physical properties of two fluvents in Turkey. Soil Tillage Res. 76, 1–11.
Krasilnikov, P., 2008. Variography of discrete soil properties. In: Krasilnikov, P., Carré, F., Montanarella, L. (Eds.),Soil Geography and Geostatistics. Concepts and Applications. JRC Scientific and Technical Reports. Office for Official Publications of the European Communities, Luxembourg, (Chapter 2), pp. 12–25.
Kravchenko, A.N., Boast Ch, W., Bullock, D.G., 1999. Multifractal analysis of soil spatial variability. Agron. J. 91, 1033–1041.
Kuzyakova, I.F., Romanenkov, V.A., Kuzyakov Ya, V., 2001. Geostatistics in soil agrochemical studies. Eurasian Soil Sci. 34 (9), 1011–1017.
Martín, M.A., Montero, E., 2002. Laser diffraction and multifractal analysis for the characterization of dry soil volume-size distributions. Soil Tillage Res. 64, 113–123.
McNeill, J.D., 1986. Geonics EM-38 Ground Conductivity Meter. Technical Note TN-21. Geonics Ltd., Canada.
Millán, H., 2004a. A pore-solid fractal approach to the fragmentation of soil materials. In: Kruhl, J.H., et, al. (Eds.), Fractals and Dynamic Systems in Geosciences. Selden and Tamm, Garching, Germany, pp. 59–65.
Millán, H., 2004b. Fragmentation of soil initiators: application of the pore-solid fractal model. Fractals 12, 357–363.
Millán, H., García-Fornaris, I., González-Posada, M., 2009. Nonlinear spatial series analysis from unidirectional transects of soil physical properties. Catena 77, 56–64.
Minasny, B., McBratney, A.B., Whelan, B.M., 2002. VESPER Version 1.6. Australian Centre for Precision Agriculture, The University of Sydney, NSW.
Pérez, L., Millán, H., González-Posada, M., 2010. Spatial complexity of soil plow layer penetrometer resistance as influenced by sugarcane harvesting: a prefractal approach. Soil Tillage Res. 110, 77–86.
Perfect, E., Díaz-Zorita, M., Grove, J.H., 2002. A prefractal model for predicting soil fragmeny mass-size distributions. Soil Tillage Res. 64, 79–90.
Perrier, E., Bird, N., 2002. Modelling soil fragmentation: the pore solid fractal approach. Soil Tillage Res. 64, 91–99.
Santra, P., Chopra, U.K., Chakraborty, D., 2008. Spatial variability of soil properties and its application in predicting surface map of hydraulic parameters in an agricultural farm. Curr. Sci. 95, 937–945.
Slavich, G.P., 1990. Determining ECa-depth profiles from electromagnetic induction measurements. Aust. J. Soil Res. 28, 443–452.
Soil Survey Staff, 2003. Keys to Soil Taxonomy, 9th edition. USDA-NRCS, Washington, DC, 332 pp.
Stat. Soft. Inc., 2003. STATISTICA (Data Analysis Software System), Version 6. Tulsa, OK.
Voltz, M., Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. J. Soil Sci. 41, 473–490.
Webster, R., 1977. Spectral analysis of gilgai soil. Aust. J. Soil Res. 15, 191–204.
Webster, R., 2008. Soil science and geostatistics. In: Krasilnikov, P., Carré, F., Montanarella, L. (Eds.),Soil Geography and Geostatistics. Concepts and Applications. JRC Scientific and Technical Reports. Office for Official Publications of the European Communities, Luxembourg, (Chapter 1), pp. 1–11.